The effect of noise on parameter estimation in systems with complicated dynamics

Justin Goodwin

Department of Mathematics and Computer Science, University of Puget Sound, Tacoma, WA 98414

Reggie Brown

Department of Physics and Department of Applied Science, College of William and Mary, Williamsburg, VA 23187-8795

Lutz Junge

Drittes Physikalisches Institut, Universität Göttingen, Bürgerstrasse 42-44, D-37073 Göttingen, Germany (February 5, 2008)

Changes in parameters of a physical device can eventually give way to catastrophic failure. This paper discusses a parameter estimation method based on synchronization between a model and time series data. In particular, we examine the robustness of the method to additive noise in the data.

I. INTRODUCTION

Scientists and engineers are often faced with the task of estimating the parameters of a physical device using time series data obtained from the device, and a mathematical model for the dynamics of the device. Because the dynamics of a model is governed by its parameters, parameter estimation amounts to adjusting the parameters in the model until its dynamics accurately mimics those of the device. Then, the parameter values of the device are assumed to be identical to those of the model.

Parameter estimation is motivated, in part, by the fact that the dynamics of all physical devices change over time. If the change arises from normal "wear and tear" then it is typically a slow change which eventually gives way to a rapid catastrophic failure. Parameter estimation aids in avoiding catastrophic failure in two ways. First, one can determine regions of parameter space associated with failure by estimating and tracking parameter values as devices are driven to failure. Second, by estimating parameter values of a device at regular intervals (and tracking these values) one can determine if and/or when the device approaches failure.

Currently, most devices in the field operate in regimes where their dynamics are regular (i.e. periodic or quasiperiodic). For these situations, slow changes in parameter values can be detected by observing shifts in the frequencies and amplitudes of Fourier spectra from time series measurements. Recently, there has been speculation about devices whose normal operating regimes have complicated and/or chaotic dynamics. Examples of such devices are bearings under high stress, electronic components with nonlinear response functions, mixing and electrochemical reactions, and high speed cutting tools. For these examples, complicated dynamics are unavoidable, and may even be desirable. More importantly, Fourier spectra will not exhibit sharp peaks. Therefore, other

methods for estimating and tracking changes in, parameter values are needed.

This paper discusses parameter estimation for systems whose underlying dynamics are complicated and/or chaotic. An early example of parameter estimation in the nonlinear dynamics literature is Hudson *et.al.* [1]. In this work the authors constructed neural net models that mimicked a few of the period doubling bifurcations found in an electrochemical experiment.

Recently, Baker et.al. [2] proposed a least squares approach that requires time series vectors which are simultaneous measurements of all of the phase space variables. The approach also requires numerical approximation of derivatives from the time series. Horbelt et.al. [3] (building on work by Baake et.al. [4]) implemented a shooting method. They used the method, and experimentally obtained data, to estimate parameters for a model of calcium release in muscle cells. Also, a novel method using symbolic dynamics has been developed for systems that exhibit complicated temporal or spatio-temporal dynamics [5].

The approach discussed in this paper estimates the parameters of a physical device by selecting values which yield "the best synchronization" between a time series from the device and a mathematical model for the device. The method is related to Kalman filtering [6] and relies on two ideas. First, if two dynamical systems are identical then it is possible for them to synchronize (i.e., both systems follow identical phase space trajectories). Second, if this behavior is stable then (for the noise free case) deviations from synchronous behavior converge to zero as differences in parameter values between the two systems converge to zero. These ideas imply that (in the absence of modeling errors) the best estimate for the parameter values of a device are those of the model which best synchronizes to a time series from the device.

The synchronization approach to parameter estimation has been discussed by Parlitz et.al. [7,8] as well as Maybhate and Amritkar [9]. The work discussed in this paper focuses on how noise in experimental data effects the accuracy of parameter estimation. Refs. [2,7–9] either did not discuss noise, or only examined small noise levels.

Furthermore, our implementation differs from the ones in previous papers. For example, Refs. [3,8,9] used vari-

ous down hill search methods to minimize the cost function. However, it is known [5,10] (and we have observed in the numerical experiments discussed below) that a cost function can have many local minima where simple down hill methods can be trapped. We avoid this problem in some of our work by using annealing to search the parameter space.

Another important difference between our approach and Refs. [7] and [9] concerns the time series used in the estimation. The previous authors assume the experimental measurements are continuously recorded. Thus, the time series has measurements at every value of time. When performing numerical experiments with this type of time series it is reasonable to replace the time series by a set of ODE's.

In our work we assume the time series consists of experimental measurements obtained at discrete increments of time (given by a sampling interval). When performing numerical experiments with this type of time series one can not replace the time series by a set of ODE's. Rather, one must use an approach that explicitly accounts for the discreteness of the time series.

Furthermore, Ref. [9] assumes that one of the phase space variables can be measured. We only assume a scalar time series corresponding to some experimental measurement. It does *not* have to be one of the phase space variables. Finally, our numerical experiments simultaneously estimate several parameters, while Ref. [9] typically estimates only one or two parameters for a three dimensional model. We have tested our methods on a high dimensional systems (an 11 dimensional generalized Rössler model).

Although modeling errors are inevitable in any mathematical description of a device, this paper does not address this issues. We conjecture that noise is the dominant obstacle associated with parameter estimation. Our examples represent electronic circuits, chemical reactions, and a high dimensional version of the Rössler system. The results indicate that this approach to parameter estimation holds promise for real applications.

The remainder of this paper is organized as follows. In Section II we define and discuss synchronization between two or more dynamical systems. In Section III we discuss parameter estimation via synchronization. Examples and the effects of noise are given in Section IV. Conclusions are presented in Section V.

II. SYNCHRONIZATION

In this section we define and discuss synchronization between two coupled deterministic dynamical systems. The numerical experiments use two different synchronization techniques, each of which is discussed. The functional forms of the uncoupled systems are identical, so they only differ in the values of their respective parameters. The coupling between the systems is denoted by

E(x,y), where $x,y \in \mathbb{R}^d$ are the instantaneous states of the two systems. The equations of motion for this arrangement are

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}; \mathbf{p}^*)$$

$$\frac{d\mathbf{y}}{dt} = \mathbf{F}(\mathbf{y}; \mathbf{p}) + \mathbf{E}(\mathbf{x}, \mathbf{y}).$$
(2)

$$\frac{d\mathbf{y}}{dt} = \mathbf{F}(\mathbf{y}; \mathbf{p}) + \mathbf{E}(\mathbf{x}, \mathbf{y}). \tag{2}$$

Equation (1), the drive system, represents the physical device, while Eq. (2), the response system, represents a mathematical model of the device. The parameters of the device and model are represented by $p^*, p \in \mathbb{R}^n$, respectively. The coupling shown in Eqs. (1) and (2) is called drive-response coupling, and is natural for our application. If the coupling in Eq. (2) vanishes when x = y (i.e., E(x, x) = 0) then synchronization is defined as $x(t_*) = y(t_*)$ at some time $-\infty < t_* < \infty$. If this occurs then, in the absence of noise, determinism implies that $\boldsymbol{x}(t) = \boldsymbol{y}(t)$ for $t > t_*$.

However, the systems we examine are chaotic, no model of a physical device can ever be exact, the state of a model can never exactly match the state of a device, and noise is never absent from real data. Therefore, synchronization between uncoupled chaotic systems never really exists because approximation synchronization $(x \simeq y)$ between chaotic systems is always unstable.

These remarks imply that the coupling between Eqs. (1) and (2) is necessary to stabilize synchronous motion. Over the last several years it has become increasingly clear that if $p = p^*$ then there are many choices of E(x,y) which will result in stable synchronization [11]. For these types of coupling, if $\|x-y\|$ is small then $\lim_{t\to\infty} \|\boldsymbol{x}-\boldsymbol{y}\| = 0$, where $\|\bullet\|$ denotes the Euclidean norm.

One form of coupling that satisfies E(x,x)=0 and often yields stable synchronization is

$$E(x, y) = K[h(x) - h(y)], \tag{3}$$

where K (the so called gain vector) could have only one nonzero component. Thus, the coupling between the drive and response systems could involve only one component of F(y; p).

It is usually impossible to simultaneously measure all of the state variables of a physical device. Typically, one experimentally measures a scalar function of the state of the device. To make contact with this more realistic case we introduce the measurement functions, h(x) and h(y). These functions represent the experimental measurement process which transforms the instantaneous state of the device or model into a scalar measurement.

The time series, $h[x(t_n)]$, represents scalar data measured at discrete times $t_n = n\Delta t$, n = 1, 2, ..., where Δt is the sampling interval. In contrast, numerical integration of the model requires E at arbitrary values of t. To overcome this problem we have used two different methods for calculating E at times not equal to one of the t_n 's. One method selects all data within a time window centered at t. Least squares is used to fit the data points within this window to a polynomial, and the polynomial is used to approximate h[x(t)] [12]. For this method the response system, Eq. (2), is continuously forced by the time series.

The second method is called sporadic driving [13]. Here, E(x,y) = 0 if t is not equal to one of the t_n 's. If t is equal to one of the t_n 's then E is given by Eq. (3). For this method, the response system (the model) is discontinuously kicked by the time series at multiples of the sampling interval, and evolves freely between kicks.

Although the two methods differ in their implementation, each leads to stable synchronization. Furthermore, each is a natural choice when confronted with a drive system that is a time series instead of ODE's.

III. PARAMETER ESTIMATION

Assume we are given a time series, $h[x(t_n)]$ for n = 1, 2, ..., N, representing experimentally measured data. Our task is to select values for the model parameters such that its dynamics are the same as those of the device which generated the data. If the parameter values of the device are p^* , and the estimated values are \hat{p} , then the goal of parameter estimation is to achieve $\hat{p} = p^*$.

The synchronization method uses coupling that guarantees stable synchronization when $p = p^*$, and the drive system is an ODE. This amounts to determining K in Eq. (3). To find \hat{p} we choose parameters, p, which minimize

$$\chi(\mathbf{p}, \mathbf{p}^*) = \left[\frac{1}{M} \sum_{n=1}^{M} (h[\mathbf{x}(n)] - h[\mathbf{y}(n)])^2\right]^{1/2}.$$
 (4)

Equation (4) explicitly notes that deviations between measured output from the model and measured output from the devices are functions of the parameters of the model, \boldsymbol{p} , and the device, \boldsymbol{p}^* . This dependence arises from the dependence of the trajectories, \boldsymbol{x} and \boldsymbol{y} , on the parameters in Eqs. (1) and (2). We denote the value of \boldsymbol{p} which minimizes Eq. (4) by $\hat{\boldsymbol{p}}$, and note that $\chi(\boldsymbol{p}^*,\boldsymbol{p}^*)=0$ in the absence of noise.

The numerical experiments typically use $M/N \sim 1/3$ or 2/3. The remaining N-M points are treated as a transient phase used to initialize synchronization.

The first two examples used annealing to search parameter space and avoid becoming trapped in local minimum far from p^* [14]. If \hat{p} is from such a local minimum then we expect the dynamics of the model to be very different from that of the device. To detect this, we performed a simple post-minimization test. After finding \hat{p} we generated a time series, h(y), from the model and used time delay embedding of h(x) and h(y) to reconstruct attractors for the data and the model [15,16].

To test the similarity of the attractors, we selected a point on the attractor reconstructed from the model time series, and calculated the distance between it and its nearest neighbor in the attractor reconstructed from the experimental data. If the dynamics of the model and the device are the same then every point on the attractor reconstructed from the model time series will have a near neighbor on the attractor reconstructed from the experimental data. More importantly, the distance to the nearest neighbor will not be large.

However, if the dynamics of the model and the device are different then some points on the model attractor will be far away from *all* points on the experimental attractor, and the distance to the nearest neighbor will be large. If this is the case then we choose a new initial condition in parameter space and minimized Eq. (4) again.

The last example used a gradient descent method (POWELL in Ref. [14]) to search the parameter space [8]. To avoid a local minima, we change each parameter by an amount of order $\sim 10^{-3}$, and the descent is started again. This procedure is repeated until we find a value for \hat{p} that is stable against such perturbations.

Numerical experiments indicate that we typically find a value of \hat{p} that is near p^* using only one or two initial conditions in parameter space. (The annealing procedure we use is probably not the best implementation of annealing. We believe that a more sophisticated implementation would remove this problem [17].) For the passivation example discussed below, the gradient descent method frequently fails to find the global minimum on the first try.

The numerical experiments use the following coupling between the drive and response systems,

$$E(x, y) = K[(h(x) + \eta) - h(y)].$$
 (5)

Here, η represents noise, and $(h(\boldsymbol{x}) + \eta)$ represents an experimental measurement that has been contaminated with additive noise. In the numerical experiments $\eta = \sigma \Sigma N(0,1)$, where N(0,1) denotes random numbers normally distributed with mean zero and unit standard deviation. Σ represents the size of the experimental signal, and was equated to the standard deviation of the clean time series, $h(\boldsymbol{x})$.

We could have examined dynamic noise. However, dynamic noise is equivalent to modeling errors which change instantaneously in time. Since we have chosen not to address modeling errors it is reasonable to ignore this type of noise.

IV. EXAMPLES

We now discuss the effect of noise on the accuracy of parameter estimation. The first example used a model of the electronic circuit shown in Fig. 1. This circuit consist of a nonlinear amplifier, N, which transforms input voltage, x, into output, $\alpha f(x)$ [18]. The parameter, α , characterizes the gain around x=0. The amplifier has linear feedback consisting of a series connection to a low-pass

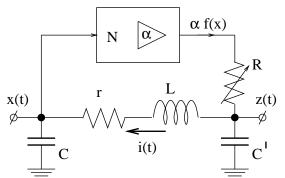


FIG. 1. A block diagram for an electronic circuit modeled by Eqs. (6) and (7).

filter, RC', and a resonant circuit LC. The system has been studied experimentally, is known to exhibit chaotic dynamics, and has been shown to synchronize with a copy of itself for many coupling schemes [18].

It is known that a good model for this circuit is the following system of ODE's

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = -x_1 - \delta x_2 + x_3$$

$$\frac{dx_2}{dt} = \gamma [F(x_1) - x_3] - \beta x_2$$
(6)

with F(x) defined by

$$F(x) = \begin{cases} 0.528\alpha & x < -1.2, \\ \alpha x (1 - x^2) & -1.2 < x < 1.2, \\ -0.528\alpha & x > 1.2 \end{cases}$$
 (7)

The numerical experiments use Eqs. (6) and (7) for the response system, continuous coupling given by Eq. (5), and parameter values $\mathbf{p}^* = [\delta, \gamma, \beta, \alpha]$ with $\delta = 0.43$, $\gamma = 0.1$, $\beta = 0.72$, and $\alpha = 16$. For this circuit it is straightforward to measure the voltages associated with \mathbf{x} . However, to make contact with other experimental devices we used the following (arbitrarily selected) measurement function

$$h(\mathbf{x}) = 2x_1 + 3x_2 + 0.5x_3. \tag{8}$$

The gain vector is K = [0, 0, 10]. In the absence of noise this coupling leads to stable synchronization when the drive and response systems are ODE's. The sampling interval is $\Delta t = 0.5$ and the data is shown in Fig. 2(a). For this example, σ , the size noise ranged from a low of 0.001 (representative of very clean data) to a high of 1 (representative of very noisy data).

We use an Adams-Bashforth-Moulton code numerically integrate the model [19]. The initial guess for p in the annealing procedure is selected at random from within the range $p_1 \in [0,1]$, $p_2 \in [0,2]$, $p_3 \in [0,5]$ and $p_4 \in [0,20]$. In the absence of noise, the method correctly estimated $\hat{p} = p^*$ to within machine precision.

The numerical experiments with noise examined 100 different value of σ . For each value we calculate $\hat{\boldsymbol{p}}$. The results of the numerical experiments (see Fig. 2(b)) indicate that $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\| \propto \sigma \|\boldsymbol{p}^*\|$ over much of the range of noise strengths. This result confirms earlier theoretical predictions [20]).

The next example uses a model for metal passivation. Passivation is a loss of chemical reactivity associated with metal corrosion in aqueous solutions. It is an electrochemical reaction that exhibits the full zoo of nonlinear dynamical behavior, including periodic cycles, period doubling, multistability, hysteresis, chaos. These behaviors are not unusual in electrochemical reactions [21–23].

We use the following passivation model [24]

$$\frac{dY}{dt} = a\theta_{\rm M} - bY$$

$$\frac{d\theta_{\rm OH}}{dt} = Y\theta_{\rm M} - [r + \exp(-\beta\theta_{\rm OH})]\theta_{\rm OH} + 2s\theta_{\rm O}\theta_{\rm M} \qquad (9)$$

$$\frac{d\theta_{\rm O}}{dt} = r\theta_{\rm OH} - s\theta_{\rm O}\theta_{\rm M},$$

where $\theta_{\rm M} \equiv (1-\theta_{\rm O}-\theta_{\rm OH})$, and the parameters $\boldsymbol{p}^* = [a,b,r,s]$ are $a=2\times 10^{-4},\ b=10^{-3},\ r=2\times 10^{-5},$ and $s=9.7\times 10^{-5}.$ We fixed the last parameter at $\beta=5$. This example is subtle because of the small size of the parameters and the attractor. To overcome this we rescale the variables by subtracting the mean and dividing by a constant roughly equal the width of the attractor in that coordinate. Time is also rescaled by a factor of 10,000 [25]. The coupling between drive and response system uses

$$h(\mathbf{x}) = 0.2x_1 + x_2 + 0.5x_3. \tag{10}$$

with gain vector $\mathbf{K} = [0, 100, 0]$. Once again $h(\mathbf{x})$ is an arbitrarily chosen function, and in the absence of noise this coupling leads to stable synchronization when the drive and response systems are ODE's. The sampling interval of the numerical experiments is $\Delta t = 0.01$ (in rescaled time) and the data is shown in Fig. 3(a).

For this example, σ , the size noise ranged from a low of 0.001 (representative of very clean data) to a high of 1 (representative of very noisy data). 100 different value of σ were examine. The results, shown in Fig. 3 (b), indicate that the relative accuracy of parameter estimation is two digits or more up to $\sigma \sim 0.1$ and about 30 % for 0 dB noise. This surprisingly robust result implies that synchronization is worthy of further investigation as a parameter estimation technique.

The final example is a high dimensional generalization of the Rössler system [26]

$$\frac{dx_1}{dt} = -x_2 + a x_1
\frac{dx_j}{dt} = x_{j-1} - x_{j+1} (j = 2, ..., L - 1)
\frac{dx_L}{dt} = b + d x_L (x_{L-1} - c).$$

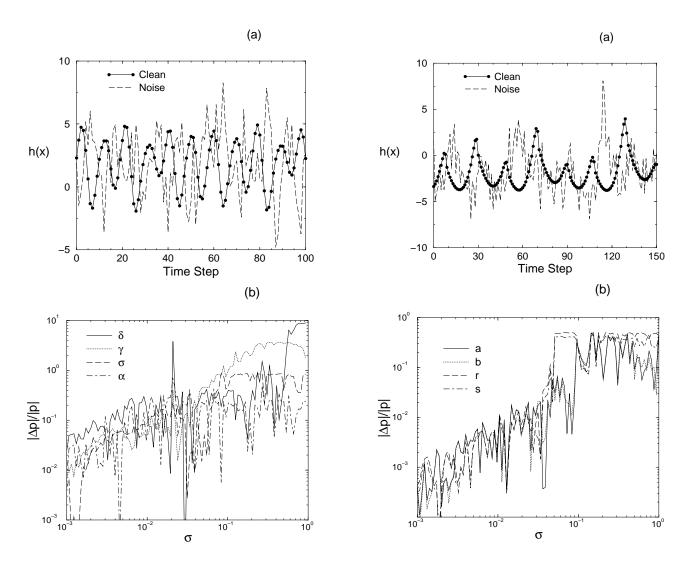


FIG. 2. The electronic circuit model. (a) Short samples of the time series used in the numerical experiments. The noisy time series represents 0 dB noise. (b) The relationship between the relative error in the parameter estimates and the size of the noise. Here, $\Delta p = \hat{p} - p^*$.

FIG. 3. The metal passivation model. (a) Short samples of the time series used in the numerical experiments. The noisy time series represents 0 dB noise. (b) The relationship between the relative error in the parameter estimates and the size of the noise. Here, $\Delta p = \hat{p} - p^*$.

Here, $p^* = [a, b, c, d]$ with a = 0.29, b = 0.1, c = 2,and d=4. This system is the well known Rössler system with an additional L-3 degrees of freedom. It has a stable chaotic attractor when L is odd. Our numerical experiments used L = 11, yielding a hyperchaotic attractor with a Lyapunov dimension of $D_L \approx 10.1$. For this example we used for synchronization the method of sporadic driving with $h(x) = x_1$ as the drive signal and a sampling interval of $\Delta t = 1$. The results were calculated for 15 different noise levels ranging from $\sigma = 10^{-5}$ to $\sigma = 0.63$. The results are shown in Fig. 4. As in the other examples, we observe a linear scaling of the parameter relative error $\frac{\Delta p}{p}$ with the noise strength σ and find good estimates up to $\sigma \sim 0.15$ The quality of the parameter estimates in relation to the noise strength is almost the same as for the electronic circuit example. This result confirms that the method can be applied to highdimensional systems, even if only very noisy time series data is available.

V. CONCLUSION

In this paper we have examined the effects of noise on the accuracy of parameter estimation for physical devices. The parameter estimation method we examined uses experimentally measured time series and a mathematical model of the device. The estimates returned by the method are those which yields the smallest deviation from synchronization between the dynamics of the model and the measured time series.

Numerical experiments indicate that this method is robust to additive noise in the time series. We find that even at 0 dB signal to noise ratio we can still obtain reasonable accuracy in the parameter estimates. The method works also for noisy time series coming from high-dimensional systems. The obvious next step is to test this method "at sea", by using experimental data coming from electronic [8], physical, chemical, ... systems.

- J. L. Hudson, M. Kube, R. A. Adomaitis, I. G. Kevrekidis, A. S. Lapedes, and R. M. Farber, *Chem. Eng. Science* 45, 2075 (1990).
- [2] G. L. Baker, J. P. Gollub, and J. A. Blackburn, Chaos 6, 528 (1996).
- [3] W. Horbelt, J. Timmer, and W. Melzer, Estimating parameters in nonlinear differential equations with applications to physiological data available from horbeltfdm.unifreiburg.de.
- [4] E. Baake, M. Baake, H. G. Bock, and K. M. Briggs, *Phys. Rev.* 45A, 5524 (1992).
- [5] X.-Z. Tang, E. R. Tracy, and R. Brown, *Physica* **102D** 253 (1997); C. S. Daw, M. B. Kennel, C. E. A. Finney, F.

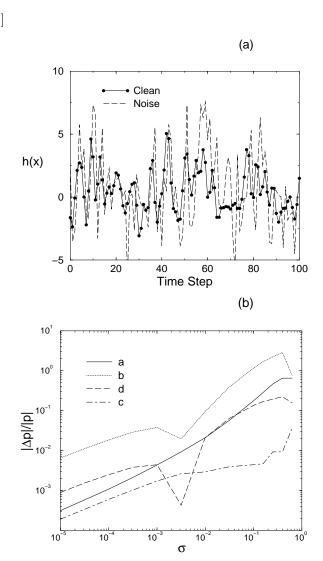


FIG. 4. The hyperchaotic generalized Rössler model. (a) Short samples of the time series used in the numerical experiments. The noisy time series represents 0 dB noise. (b) The relationship between the relative error in the parameter estimates and the size of the noise. Here, $\Delta p = \hat{p} - p^*$.

- T. Connolly, Phys. Rev. 57E, 2811 (1998); X.–Z. Tang, E. R. Tracy, A. D. Boozer, A. deBrauw, and R. Brown, Phys. Letts. 190A, 393 (1994).
- [6] W. L. Brogan, Modern Control Theory, 3rd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1991).
- [7] U. Parlitz, Phys. Rev. Letts. 76, 1232 (1996).
- [8] U. Parlitz, L. Junge, and L. Kocarev, *Phys. Rev.* 54E, 6253 (1996).
- [9] A. Maybhate and R. E. Amritkar, Phys. Rev. 59E, 284 (1999).
- [10] S. V. Ershov, Phys. Letts. 177A, 180 (1993).
- [11] Chaos 7 No. 4 (1997) and references therein.
- [12] This code, polfit.f, can be found in SLATEC, a public domain source code library.
- [13] U. Parlitz, L. Kocarev, T. Stojanovski, and L. Junge, *Physica* **109D**, 139 (1997).
- [14] Chapter 10 of W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
- [15] K. T. Alligood, T. D. Sauer, and J. A. Yorke, CHAOS and introduction to dynamical systems, (Springer Verlag, New York, 1996).
- [16] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, Rev. Mod. Phys. 65, 1331 (1993).
- [17] L. Ingber, em Math. Comp. Modelling 8, 29 (1993).
- [18] N. F. Rulkov, A. R. Volkovskii, A. Rodriguez-Lozano, E. Del Rio, and M. G. Velarde, Int. J. of Bifur. and Chaos, 2, 669 (1992).
- [19] This code, deabm.f, can be found in SLATEC, a public domain source code library.
- [20] R. Brown, N. F. Rulkov, and N. B. Tufillaro, Phys. Rev. 49E, 3784 (1994).
- [21] J. B. Talbot and R. A. Oriani, AiChE symp. series, Electrochem. Eng. Appl. 83, 64 (1992).
- [22] D. Haim, O. Lev, L. M. Pismen, and M. Sheintuch, Chem. Eng. Science, 47, 3907 (1992).
- [23] M. T. M. Koper, J. Chem. Soc. Faraday, Trans 94, 1369 (1998).
- [24] J. Kevin McCoy, P. Parmananda, W. Rollins, and A. J. Markworth, J. Mater. Res 8, 1858 (1993).
- [25] In this paper we scaled the parameters and the variables. A more complete discussion is in R. Brown, N. F. Rulkov, and E. R. Tracy, *Phys. Rev.* 49E, 3784 (1994).
- [26] G. Baier and M. Sahle, Phys. Rev. **51E**, R2712 (1995).